



Contact Details:

Email: Soham.Ganguly@cognizant.com
Mobile: +91-8777323212
Portfolio: [Soham's Portfolio](#)

Key Competency Areas

🛠 Tools & Technologies

- **Data Analysis & Processing:** Advanced Excel, SQL, ETL, Data Cleaning, Python (Pandas, NumPy), Feature Engineering
- **Data Visualization & Storytelling:** Power BI (DAX), Excel Pivot Charts
- **Gen AI/LLM:** Transformers, Prompt , Design, Frameworks & Engineering (Zero-shot, Few-shot, Chain-of-Thought, Persona), GPT-4, Gemini, Claude, OpenAI, RAG, Multimodal LLMs, Image captioning (PIL), LangChain, HuggingFace, FAISS, Chroma DB, Agentic AI, MCP Server
- **Cloud & AI Platforms:** GCP (Google Cloud Platform), Vertex AI, Azure AI, Fine Tuning, Indexing

💼 **Core Competencies:** Data Analytics & Research, Advisory to the Offices of CXOs, Global Stakeholder Management, Business Acumen, Continuous Improvement

💻 **Industries:** E-commerce, Internet, Media & Entertainment, Management Consulting,

Education

Executive Development Program (People Analytics) 2021 -XLR/Jamshedpur

Bachelor of Engineering (Computer Science) 2012 -University of Burdwan

Profile Summary

A passionate Gen AI/LLM Engineer specialized in developing and deploying machine learning models, with a strong focus on Natural Language Processing (NLP) and Large Language Models (LLMs). Currently learning Agentic AI Solutions using MCP architecture.

Key Engagements in Cognizant or Previous Company

- Part of a team for the implementation of an Agentic AI solution using the LangChain framework for AI agent orchestration, MetaLlama-v3.2-instruct as the LLM for NLU, ChromaDB for embeddings and vectorization, memory hub and context handling to automate and improve the performance of a customer service agent which improved accuracy by 30% and reduced response rates by 50% and made necessary tool calls for escalation to a human agent in case of a complex query.
- Part of a global team responsible for training an under development Multi-Modal LLM called Nightingale for the statistics and mathematics domain which accepts inputs both as texts, images, slides and articles with relevant prompt-response pairs to make sure the best response is generated for a user-given prompt by the LLM using PIL, LangChain, FAISS and Gradio as frontend.
- Senior member of a global team of 20+ members responsible for training an under development Mathematical translation LLM called Wukong Legend which translated a complex Mathematical problem from English to Bengali using Azure AI language translation libraries.

Sample Projects in Cognizant/Previous Company

1. Chatbot with Memory: A contextual Memory chatbot using the mistral-7b-instruct-v2.0 model, optimized to run on a CPU.

Technologies used: Python, LangChain, Gradio, NLP, LLM (Hugging face libraries)

GitHub link: [Click Here](#)

2. Contextual Multimodal RAG with Image Captioning: A contextual multimodal RAG model with image captioning capabilities. Accepts doc, docx, csv, png and jpg files as inputs with a Gradio Web UI frontend.

Technologies Used: Python, LangChain, Huggingface, FAISS, Gradio

Google Colab link: [Click Here](#)

3. Agentic AI Web Search Summarizer: A multi agent Agentic AI solution with a Gradio Web UI frontend which deals with:

- DuckDuckGo website search to search the relevant webpages to answer the user query.
- A text summarization tool which summarized the results obtained from the previous step output.

Technologies Used: Python, LangGraph, LangChain, Tiny Llama 1.1-B, ChromaDB, Gradio, DuckDuckGo search API

GitHub link: [Click Here](#)

Courses and Certifications	Employment History
AI For Everyone View Certificate - DeepLearning.AI	Senior Associate – Gen AI and LLM (Full-Time) May 2025 – Present <i>Cognizant</i> <ul style="list-style-type: none"> Joined on 5th May 2025. Allocated to Apple account, waiting for the project work to start, building personal projects till project work starts. Built a Data Analytics & Visualization Agentic AI Solution POC for Football Association UK client. (Click Here) Completed Nvidia Gen AI certification, Google Cloud Generative AI Leader Certification.
Python Certificate View Certificate - HackerRank	Senior Gen AI & LLM Engineer (Freelancer) Jan 2025 – Apr 2025 <i>Oneforma (Centific)</i> <ul style="list-style-type: none"> Applied mathematical concepts and terminologies to translate the given complex Math problem in a simplified Bengali language consisting optimal number of word counts using Azure AI translator and Azure AI language services. Involved in multiple Data Annotation, Summarization and Translation projects using Data and research driven mindset to ensure high quality training data provided to train the LLM model to reduce hallucination from the output. Implemented advanced RAG (Retrieval Augmented Generation) using the LangChain framework and vector databases like FAISS and ChromaDB and performed fine tuning within the LLMs. Performed Text and Image based embeddings using OpenAI, Vertex AI text embeddings and ResNet, PIL.
SQL Certificate View Certificate - HackerRank	Senior Gen AI & LLM Engineer (Freelancer) Jan 2024 – Dec 2024 <i>Soul AI (Outlier)</i> <ul style="list-style-type: none"> Applied Statistical and mathematical concepts to analyze the images, articles, graphs and charts to provide the best promptresponse pair in a simplified English language consisting of the optimal number of word counts. Used Prompt Engineering techniques and frameworks such as Zero shot, few shot, multi shot, Chain of thoughts (CoT) etc. to instruct the AI with the best prompt-response pairs for the LLM. Utilized the Vertex AI Workbench, freeform and Google Gemini 1.5 and 2 pro models to perform multimodal text, image, video and market data analysis using Python kernel on Google cloud platform and console.
Introduction to Generative AI View Certificate - Google	Team Lead – Global Insights & Data Analytics (Full-Time) Sep 2021 – Dec 2023 <i>Capgemini</i> <ul style="list-style-type: none"> Orchestrated and directed a high-performing team of 12 members, achieving a remarkable 233% growth in team size in 2022 and 30% growth in 2023, while maintaining 0% attrition and an average pulse dashboard score of 9.3 (out of 10) within the team. Designed and managed ETL processes to extract and transform HR data from multiple sources, processing over 10 million records for integration into Workday and SAP SuccessFactors HCM platforms. Implemented automated data validation checks, decreasing data errors by 22% during ETL workflows and ensuring reliable HR reporting. Used Python libraries like Pandas, NumPy for Data Analysis. Optimized complex SQL queries, reducing data processing times by 40% and enhancing performance during large-scale data transfers. Authored detailed ETL workflow and data mapping documentation, adopted as a team standard, reducing onboarding time for new members by 25%. Facilitated cross-functional workshops, enhancing team communication and accelerating project delivery timelines by 10%. Analyzed Workday Peakon Pulse surveys by implementing Sentiment Analysis using Python NLP libraries, predicted the concerned areas and suggested improvements to enhance the employee engagement experience which boosted the employee response rates from 72% to 88% and a better employee engagement experience ensuring the optimal level of automation in the overall process.
Introduction to Large Language Models View Certificate - Google	Senior Data Analyst (Full-Time) Feb 2019 – Sep 2021 <i>MSH Talent Solutions</i> <ul style="list-style-type: none"> Involved in end-to-end data analysis process starting from data collection, ETL processes, transforming the data using SQL and Power Query, apply data cleaning procedures to make the data ready for analysis ensuring data quality, defining the data models and dependencies. Applied statistical data analysis methods to analyze data, used data visualization tools like Power BI (DAX) and advanced excel (Pivot tables and Pivot charts) to make the dashboards and preparing the presentations reports using data storytelling skills to provide insights to the senior executive stakeholders and the CXOs of the company. Weekly and monthly data visualization status report generation using Advanced Excel and Power BI cloud service.
Prompt Design in Vertex AI View Certificate – Google	
Google Cloud Generative AI Leader View Certificate – Google	
NVIDIA GSI Curriculum 2025 View Certificate - Google	
Rewards & Recognitions	
<input checked="" type="checkbox"/> Best Employee (ManpowerGroup India, Q3 - 2018) Awarded Best employee in Freshworks RPO team for closing six critical positions in Q3-2018.	
<input checked="" type="checkbox"/> Best Performer (MSH Talent Solutions, for FY 2019-20 and 2020-21) Awarded Best performer for generating total revenue of USD 1.3M combined in two FY.	
<input checked="" type="checkbox"/> R&R (Capgemini, Q1, Q2, Q4 - 2022 & Q2 – 2023) Won Out of the box, Rising Star, XtraMile and Outstanding contribution in Service Delivery award for HRO Hackathon Automation idea submission, providing solutions which saved around € 820k as cost of hire for recruitment teams.	